CAPíTULO 6

Multicolinealidad

6.1. Introducción

Uno de los supuestos básicos del modelo lineal general

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

establece que las variables explicativas son linealmente independientes, es decir, la igualdad

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_k \mathbf{x}_k = 0$$

se verifica para $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$, donde $\mathbf{x}_i\,(i = 1, \ldots, k)$ es la i-ésima columna de $\mathbf{X}$ y $\lambda_i\,(i = 1, \ldots, k)$ es un número real, $\lambda_i \in \Re$. Este supuesto asegura que la matriz $\mathbf{X}$ de orden $N \times k$ tiene un rango igual a k, el número de columnas en $\mathbf{X}$. De aquí, la matrix $\mathbf{X}'\mathbf{X}$ de orden $k \times k$ tiene también un rango igual a k, su determinante $|\mathbf{X}'\mathbf{X}|$ difiere de cero y su inversa puede calcularse mediante el método de la matriz adjunta como

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{|\mathbf{X}'\mathbf{X}|} Adj(\mathbf{X}'\mathbf{X})$$

En definitiva, el supuesto de ausencia de multicolinealidad garantiza que el sistema de ecuaciones normales

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$$

es un sistema de Cramer, que siempre admite una solución única y que ésta viene dada por el estimador de mínimos cuadrados

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Ahora bien, cuando las columnas de la matriz $\mathbf{X}$ son linealmente dependientes, la igualdad

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_k \mathbf{x}_k = 0$$

se verifica para algunos $\lambda_i \in \Re\,(i = 1, \ldots, k)$ no nulos. Entonces el rango de la matriz $\mathbf{X}$ es menor que k, la matriz $\mathbf{X}'\mathbf{X}$ es singular, su determinante $|\mathbf{X}'\mathbf{X}|$ es igual a cero y el estimador de mínimos cuadrados queda indeterminado.

Definición 44. *Existe multitolinealidad exacta o perfecta cuando las columnas de la matriz* $\mathbf{X}$ *son linealmente dependientes.*

Ejemplo 10. *Las columnas de la matriz*

$$\mathbf{X} = \begin{pmatrix} 1 & 3 & 7 \\ 1 & 2 & 5 \\ 1 & 4 & 9 \\ 1 & 6 & 13 \end{pmatrix}$$

son linealmente dependientes: la tercera columna es la primera más dos veces la segunda.

A veces las columnas de la matriz $\mathbf{X}$ son casi linealmente dependientes

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_k \mathbf{x}_k \simeq 0$$

Aquí, la matriz $\mathbf{X}$ tiene un rango igual a k, la matriz $\mathbf{X}'\mathbf{X}$ es no singular y el estimador de mínimos cuadrados puede calcularse.

DEFINICIÓN 45. *Existe multitolinealidad aproximada cuando las columnas de la matriz $\mathbf{X}$ son casi linealmente dependientes.*

EJEMPLO 11. *Las columnas de la matriz*

$$\mathbf{X} = \begin{pmatrix} 1 & 3 & 7{,}01 \\ 1 & 2 & 5 \\ 1 & 4 & 9 \\ 1 & 6 & 13 \end{pmatrix}$$

son casi linealmente dependientes. El determinante de la matriz $\mathbf{X}'\mathbf{X} = 0{,}0024$.

Observación 33. *La presencia de multicolinealidad aproximada en un modelo de regresión múltiple no viola ninguno de los supuestos básicos. Se deduce entonces, por el teorema de Gauss-Markov, que bajo multicolinealidad aproximada el estimador de mínimos cuadrados es el estimador lineal e insesgado de menor varianza.*

6.2. Consecuencias de la multicolinealidad

Consideremos la ecuación de regresión múltiple

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

y supongamos que $X_{3i} = kX_{2i}$ donde k es un número real conocido. Esta ecuación presenta multicolinealidad exacta y no es posible obtener los estimadores de mínimos cuadrados de los coeficientes. Sin embargo, sí es posible estimar combinaciones de parámetros. En efecto, incorporando la relación $X_{3i} = kX_{2i}$ en la ecuación de regresión

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 k X_{2i} + u_i$$

se obtiene

$$Y_i = \beta_1 + (\beta_2 + k\beta_3) X_{2i} + u_i$$

que es la regresión simple de Y_i sobre X_{2i}, y que puede estimarse por mínimos cuadrados sin problemas. La ordenada estimada es precisamente la estimación de β_1 que no puede estimarse en la regresión múltiple, y la pendiente estimada es una estimación de la combinación lineal $\beta_2 + k\beta_3$. Ahora bien, no es posible obtener estimaciones individuales de β_2 y β_3. De aquí, la principal consecuencia de la multicolinealidad exacta es que no podemos medir los efectos individuales de las variables explicativas. En otras palabras, no podemos estimar la respuesta de Y ante un cambio unitario de X_2, porque al cambiar X_2 también lo hace X_3. Solo podemos estimar la respuesta de Y ante un cambio unitario de X_2 y un cambio igual a k de X_3, respuesta que es igual a $\beta_2 + k\beta_3$.

PROPOSICIÓN 58. *Cuando dos o más variables explicativas guardan una relación exacta entre ellas, no es posible medir los efectos individuales de estas variables sobre la variable dependiente.*

Observación 34. Si la combinación lineal estimada $\beta_2 + 2\beta_3$ es, por ejemplo, igual a 0.8, y por un estudio previo o alternativo se dispone de una estimación del parámetro $\hat{\beta}_3$, digamos 1.2, entonces es posible estimar el parámetro β_2 como $\hat{\beta}_2 = 0{,}8 - 2 \times 1{,}2$.

La regresión simple estimada $Y_i = \hat{\beta}_1 + \hat{\beta}_2^* X_{2i} + \hat{u}_i$ ($\hat{\beta}_2^* \beta_2 + k\beta_3$) puede usarse también para predecir los valores de Y_i a partir de los valores conocidos de X_{2i}. En este sentido, se dice que la multicolinealidad exacta no es un problema para la predicción.

Para ver las consecuencias de la multicolinealidad aproximada hacemos uso del modelo de regresión particionado

$$\mathbf{y} = \beta_i \mathbf{x}_i + \mathbf{X}_s \boldsymbol{\beta}_s + \mathbf{u}$$

donde $\mathbf{x}_i$ es la i-ésima columna de la matriz $\mathbf{X}$ y $\mathbf{X}_s$ son las otras columnas de $\mathbf{X}$. El estimador de mínimos cuadrados del coeficiente β_i es

$$\hat{\beta}_i = (\mathbf{x}_i' \mathbf{M}_s \mathbf{x}_i)^{-1} \mathbf{x}_i' \mathbf{M}_s \mathbf{y}$$

y su varianza

$$V(\hat{\beta}_i) = \sigma^2 (\mathbf{x}_i' \mathbf{M}_s \mathbf{x}_i)^{-1}$$

Definición 46. *La precisión o acuracidad de un estimador insesgado se define en relación a la inversa de su varianza. Cuanto menor sea la varianza de un estimador, tanto mayor será la precisión o acuracidad del mismo.*

Proposición 59. *La varianza de $\hat{\beta}_i$ puede expresarse como*

$$V(\hat{\beta}_i) = V(\hat{b}_i) \frac{1}{1 - R_i^2} \tag{6.1}$$

en donde $V(\hat{b}_i)$ es la varianza de la estimación de la pendiente en la regresión simple de Y sobre X_i, y R_i^2 es el coeficiente de determinación en la regresión múltiple de X_i sobre las restantes variables explicativas.

Demostración. La varianza de la pendiente estimada en regresión simple es

$$V(\hat{b}_i) = \frac{\sigma^2}{\sum_{t=1}^N (X_{it} - \bar{X}_i)^2}$$

Por otro lado, la forma cuadrática $\mathbf{x}_i' \mathbf{M}_s \mathbf{x}_i$ se interpreta como la suma de cuadrados de los residuos en la regresión de X_i sobre las restantes variables explicativas $\mathbf{X}_s$. El cociente de esta suma de cuadrados de los residuos y la suma de cuadrados total para la variable X_i es

$$\frac{\mathbf{x}_i' \mathbf{M}_s \mathbf{x}_i}{\sum_{t=1}^n (X_{it} - \bar{X})^2} = 1 - R_i^2$$

De aquí,

$$V(\hat{\beta}_i) = \sigma^2 (\mathbf{x}_i' \mathbf{M}_s \mathbf{x}_i)^{-1} = \frac{\sigma^2}{\sum_{t=1}^n (X_{it} - \bar{X})^2 (1 - R_i^2)} = V(\hat{b}_i) \frac{1}{1 - R_i^2}$$

□

La ecuación (6.1) muestra que la varianza de $\hat{\beta}_i$ aumenta con R_i^2. Esta varianza toma el valor más pequeño, $V(\hat{b}_i)$, cuando las variables explicativas son ortogonales, $R_i^2 = 0$, y el valor máximo cuando existe multicolinealidad exacta, $R_i^2 = 1$. Entre estos dos extremos, podemos graduar la multicolinealidad como baja, media y alta.

Definición 47. *El factor incremento de varianza se define como el cociente de la varianza de la estimación de una pendiente en regresión múltiple y la varianza de la misma pendiente en regresión simple*

$$FIV_i = \frac{V(\hat{\beta}_i)}{V(\hat{b}_i)} = \frac{1}{1 - R_i^2}$$

La principal consecuencia de una multicolinealidad alta, $R_i^2 \simeq 1$, es que las varianzas de las estimaciones asociadas a las variables colineales son muy grandes. Ahora bien, aún siendo estas varianzas muy grandes, son las menores que puede alcanzar un estimador lineal e insesgado.

Proposición 60. *En un modelo de regresión multiple con alta multicolinealidad los contrastes t están sesgados hacia la aceptación de la hipótesis nula de no significación individual, $H_0 : \beta_i = 0$.*

Demostración. El estadístico t para la estimación $\hat{\beta}_i$ es

$$t_i = \frac{\hat{\beta}_i}{\sqrt{V(\hat{\beta}_i)}} = \frac{(\mathbf{x}_i'\mathbf{M}_s\mathbf{x}_i)^{-1}\mathbf{x}_i'\mathbf{M}_s\mathbf{y}}{\hat{\sigma}\sqrt{(\mathbf{x}_i'\mathbf{M}_s\mathbf{x}_i)^{-1}}} = \frac{\mathbf{x}_i'\mathbf{M}_s\mathbf{y}}{\hat{\sigma}\sqrt{\mathbf{x}_i'\mathbf{M}_s\mathbf{x}_i}}$$

Cuando $R_i^2 \to 1$, $\mathbf{x}_i'\mathbf{M}_s\mathbf{x}_i \to \infty$ y la ratio $t_i \to 0$. □

En resumen, las consecuencias de la multicolinealidad aproximada son

1. Las varianzas de las estimaciones son muy grandes.
2. Las ratios t son pequeñas y están sesgadas a la aceptación de la hipótesis de no significación individual.
3. Los intervalos de confianza, $\hat{\beta}_i \pm c\sqrt{V(\hat{\beta}_i)}$, son muy amplios, incluyendo valores negativos y positivos.

6.3. Detección de la multicolinealidad

La multicolinealidad exacta siempre se detecta al tratar de estimar el modelo, aunque no seamos conscientes de su presencia: el determinante de la matriz $\mathbf{X}'\mathbf{X}$ es nulo y no es posible calcular su inversa. En cambio, la detección de la multicolinealidad alta resulta en la práctica más problemática.

Parece natural usar el determinante de la matriz $\mathbf{X}'\mathbf{X}$ como primer método de detección de multicolinealidad. Si el valor de este determinante es aproximadamente cero, entonces podemos pensar que existe multicolinealidad alta. El problema es que el valor de este determinante depende de las unidades de medida de las variables explicativas. Así, pueden obtenerse conclusiones diferentes dependiendo de si los datos se expresan en unidades o en millones de unidades.

Este inconveniente puede evitarse utilizando datos tipificados. La variable X_i tipificada es

$$X_{it}^* = \frac{X_{it} - \bar{X}_i}{\sqrt{\sum_{t=1}^{N}(X_{it} - \bar{X}_i)^2/n}}$$

Si definimos $\mathbf{X}_*$ como la matriz que contiene los datos tipificados de las variables explicativas, entonces $\mathbf{X}'_*\mathbf{X}_*$ es la matriz $(k-1)\times(k-1)$ de correlaciones simples

$$\mathbf{R} = \begin{pmatrix} 1 & r_{23} & \dots & r_{2k} \\ r_{32} & 1 & \dots & r_{3k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k2} & r_{k3} & \dots & 1 \end{pmatrix}$$

en donde

$$r_{ij} = \frac{\sum_{t=1}^{N} X^*_{it} X^*_{jt}}{n} = \frac{\sum_{t=1}^{N}(X_{it}-\bar{X}_i)(X_{jt}-\bar{X}_j)}{\sqrt{\sum_{t=1}^{N}(X_{it}-\bar{X}_i)^2 \sum_{t=1}^{N}(X_{jt}-\bar{X}_j)^2}}$$

El determinante de $\mathbf{R}$ no depende de unidades de medida y toma valores entre 0 y 1. $|\mathbf{R}|$ será igual a 1 cuando todas las correlaciones simples entre pares de variables explicativas sean iguales a cero, $r_{ij}=0$, y $|\mathbf{R}|$ será igual a cero cuando al menos una correlación simple sea igual a 1, es decir, cuando dos columnas de la matriz $\mathbf{X}_*$ sean linealmente dependientes.

La existencia de correlaciones simples altas, mayores que 0.8 ó 0.9, entre variables explicativas es también una señal de multicolinealidad alta. Sin embargo, la existencia de correlaciones simples bajas, menores que 0.8 ó 0.9, entre variables explicativas no es una señal de multicolinealidad baja. Es posible que tres variables explicativas X_{2t}, X_{3t} y X_{4t} estén exactamente relacionadas, $X_{4t} = X_{2t} + X_{3t}$, y que las correlaciones simples r_{23}, r_{24} y r_{34} sean pequeñas. De aquí, una mejor medida de multicolinealidad sería el coeficiente de correlación múltiple en la regresión auxiliar de X_i sobre el resto de variables explicativas, R_i (la raíz cuadrada del coeficiente de determinación, R_i^2).

Otras medidas de multicolinealidad basadas en R_i^2 son el factor incremento de varianza

$$FIV_i = \frac{1}{1-R_i^2}$$

y su inverso, el factor tolerancia de varianza $TOL_i = 1-R_i^2$. Si $FIV_i > 10$ o $TOL_i < 0{,}1$, entonces la variable X_i está altamente relacionada con el resto de explicativas. Klein sugiere comparar R_i^2 con el coeficiente de determinación en la ecuación de interés: la multicolinealidad sería un problema si $R^2_{1,2\ldots k} < R_i^2$.

Un estadístico F alto y varios estadísticos t bajo también se considera un síntoma de multicolinealidad alta. En esta situación, con el estadístico F se rechaza la hipótesis de no significación global $H_0 : \beta_2 = \cdots = \beta_k = 0$, mientras que con los estadísticos t se aceptan las hipótesis de no significación individual $H_0 : \beta_i = 0$. Los estadísticos t son bajos porque las varianzas de las estimaciones $\hat{\beta}_i$ son muy grandes; el estadístico F es alto si lo es el R^2 de la regresión:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

La medida más satisfactoria de multicolinealidad se basa en los autovalores de la matriz $\mathbf{X}'\mathbf{X}$. Recordemos que una matriz simétrica y definida positiva puede escribirse como

$$\mathbf{X}'\mathbf{X} = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}'$$

donde $\mathbf{C}$ es una matriz ortogonal, $\mathbf{C}' = \mathbf{C}^{-1}$, y $\mathbf{\Lambda}$ es una matriz diagonal que contiene los autovalores, $\mathbf{\Lambda} = diag\{\lambda_1, \ldots, \lambda_k\}$. Las columnas de $\mathbf{C}$ contienen los autovectores asociados a los autovalores λ_i. Se cumple que el determinante de una matriz es el producto de sus autovalores

$$|\mathbf{X}'\mathbf{X}| = |\mathbf{C}\mathbf{\Lambda}\mathbf{C}'| = |\mathbf{C}||\mathbf{\Lambda}||\mathbf{C}'| = |\mathbf{\Lambda}| = \lambda_1 \times \lambda_2 \times \cdots \times \lambda_k$$

Es claro que el determinante de $\mathbf{X}'\mathbf{X}$ será cero cuando alguno de los autovalores sea nulo. Por tanto, podemos pensar que existe multicolinealidad alta cuando alguno de los autovalores sea pequeño. El problema es que los autovalores de esta matriz también dependen de las unidades de media. Sin embargo, el cociente de dos autovalores es una magnitud adimensional.

El número de condición de la matriz $\mathbf{X}$ se define como

$$\kappa = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$$

donde λ_{max} y λ_{min} son los autovalores de $\mathbf{X}'\mathbf{X}$ más grande y más pequeño, respectivamente. Si $\kappa > 30$, entonces existe multicolinealidad alta.

6.4. Soluciones a la multicolinealidad

Una vez que hemos estimado una ecuación de regresión múltiple y hemos aplicado los métodos de detección de multicolinealidad, si descubrimos la presencia de multicolinealidad alta ¿qué debemos hacer? Tenemos las siguientes posibilidades:

1. No hacer nada. Posiblemente la mejor decisión. La multicolinealidad es una deficiencia de los datos, no de los métodos estadísticos. Por la alta relación entre las variables explicativas, no podemos medir sus efectos individuales con precisión. Sin embargo, nuestras estimaciones son las mejores estimaciones lineales e insesgadas. Vamos a ver que la mayoría de los remedios propuestos suelen ser peores que la enfermedad, no sólo no corrigen el problema de la multicolinealidad sino que además introducen otros problemas adicionales.
2. Obtener más datos. Al utilizar datos adicionales aumentamos el número de filas de la matriz $\mathbf{X}$ y esperamos que los nuevos datos no estén muy relacionados para se puede mitigar el problema. Ahora bien, no siempre es posible acceder a nuevos datos y, cuando esto es posible, no hay ninguna razón para los nuevos datos no estén tan interrelacionados con los ya usados.
3. Incorporar estimaciones de otros estudios. Imaginemos que se quiere estimar la ecuación de regresión

   $$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$$

   usando datos de series temporales, y supongamos que las variables X_{2t} y X_{3t} están fuertemente relacionadas. Una solución sería recurrir a estimaciones basadas en datos de sección cruzada

   $$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i$$

   Si sustituimos β_3 por $\hat{\alpha}_3$ entonces la ecuación a estimar sería

   $$Y_t^* = \beta_1 + \beta_2 X_{2t} + u_t$$

donde los datos de la nueva variable dependiente son $Y_t^* = Y_t - \hat{\alpha}_3 X_{3t}$

4. Eliminar variables explicativas. En el ejemplo, anterior podemos omitir una de las variables explicativas y estimar la ecuación

$$Y_t = \delta_1 + \delta_2 X_{2t} + u_t$$

El problema es que se incurre en un error de especificación por omisión de variable relevante que, como veremos en el siguiente capítulo, da lugar a estimaciones sesgadas.

5. Transformar los datos. Con datos de sección cruzada se recomienda utilizar cocientes de variables, esto es, si la ecuación de interés es

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i$$

dividiendo por X_{3i} obtenemos

$$\frac{Y_i}{X_{3i}} = \beta_1 \frac{1}{X_{3i}} + \beta_2 \frac{X_{2i}}{X_{3i}} + \beta_3 + \frac{u_i}{X_{3i}}$$

y esperamos que la relación entre $1/X_{3i}$ y X_{2i}/X_{3i} sea menor que la existente entre X_{2i} y X_{3i}. El problema con esta transformación es que se introduce heterocedasticidad en el modelo. Con datos de series temporales se recomienda utilizar datos en primeras diferencias. La primera diferencia de la variable dependiente Y_t es $\Delta Y_t = Y_t - Y_{t-1}$; y la primera diferencia de la variable explicativa X_i, $\Delta X_{it} = X_{it} - X_{it-1}$. Si la ecuación de regresión

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$$

se escribe en el instante $t-1$

$$Y_{t-1} = \beta_1 + \beta_2 X_{2t-1} + \beta_3 X_{3t-1} + u_{t-1}$$

Restando de la primera ecuación la segunda,

$$\Delta Y_t = \beta_2 \Delta X_{2t} + \beta_3 \Delta X_{3t} + e_t$$

donde el error $e_t = u_t - u_{t-1}$ presenta autocorrelación.

6. Usar el estimador cresta, que se define como

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X} + r\mathbf{I}_k)^{-1}(\mathbf{X}'\mathbf{y})$$

donde r es un número real que suele fijarse en 0.01. Este estimador es sesgado, pero puede tener una varianza menor que la del estimador de mínimos cuadrados.

7. Usar componentes principales. Se trata de un método para reducir el número de variables explicativas sin perder demasiada información relevante. Sea $\mathbf{X} = (\mathbf{i} \quad \mathbf{X}_s)$. El primer componente principal de la matriz $\mathbf{X}_s$ es la combinación lineal de las columnas de $\mathbf{X}_s$, $\mathbf{z}_1 = \mathbf{X}_s\mathbf{c}_1$ que tiene la mayor varianza. El segundo componente principal de la matriz $\mathbf{X}_s$ es la combinación lineal de las columnas de $\mathbf{X}_s$, $\mathbf{z}_2 = \mathbf{X}_s\mathbf{c}_2$, que es ortogonal a $\mathbf{z}_1$ y que tiene la mayor varianza. Así, sucesivamente. Los vectores $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{k-1}$ son los autovectores de la matriz $\mathbf{X}_s'\mathbf{M}_i\mathbf{X}_s$. De acuerdo con esta solución, habría que realizar de regresión de Y

sobre los primeros $d < (k-1)$ componentes principales. Aquí, se presenta el problema de interpretar las nuevas variables $\mathbf{z}_i$.

Resumen

1. Formas de multicolinealidad: exacta y aproximada
2. Consecuencias:
   a) Multicolineadad exacta: no es posible estimar efectos individuales.
   b) Multicolinealidad aproximada: varianzas grandes.
3. Detección:
   a) Correlaciones simples, correlaciones múltiples, número de condición.
4. Solución:
   a) No hacer nada.

Palabras clave

Multicolinealidad exacta
Multicolinealidad aproximada
Precisión del estimador
Grado de multicolinealidad
Factor incremento de varianza

Ejercicios

1. Explique de qué factores depende la precisión del estimador de mínimos cuadrados y cuál es el efecto del tamaño muestral sobre la misma.
2. Demuestre que
$$V(\hat{\beta}_i) = \sigma^2 \sum_{j=1}^{k} \frac{c_{ij}}{\lambda_j}$$
en donde c_{ij} $(j = 1, \ldots, k)$ son los elementos en la fila i de la matriz de autovectores $\mathbf{C}$.
3. En el modelo de regresión múltiple
$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i$$
se cumple que $X_{2i} = 3X_{4i}$. Indique qué parámetros son estimables (1) cuando no se dispone de información a priori sobre ningún coeficiente y (2) cuando se sabe que $\beta_4 = 2$.
4. Comente la siguiente proposición: en el modelo de regresión múltiple
$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + u_i$$
existe multicolinealidad exacta porque la segunda variable es el cuadrado de la primera. Este problema puede corregirse aplicando la transformación logarítmica y estimando la ecuación
$$\log(Y_i) = \beta_1 + \beta_2 \log(X_i) + \beta_3 \log(X_i^2) + u_i$$